Benchmarking Imitation and Reinforcement Learning for NPC players in casual video games

Gema Parreño Piqueras

gema.parreno.piqueras@gmail.com

Abstract

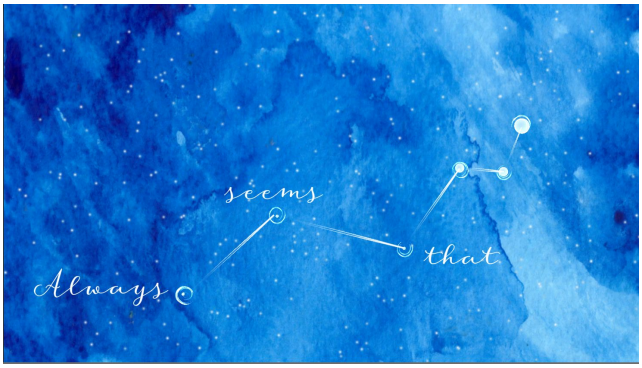
This work aims to present the principles from the game design perspective and lessons learned of two techniques adopted for solving the NPC behavior in casual language-oriented games in a Reinforcement Learning discrete and partially observable environment. It might be useful for other game developers as it offers an example of designing companionship in NPCs from the game perspective and to speed up their developments for implementing machine learning in NPC players, showing that a designed heuristic function and Imitation Learning approach can speed up developments with respect to a Reinforcement Learning approach for a deterministic output.

The video game Mempathy [1]

Mempathy is a video game narrative experience that transforms the relationship with anxiety. The video game's goal is to offer a reflective experience, and the winning state is defined by a feeling of advancement and companionship towards this mental health topic. The idea of progress is supported in art by watercolor progression and on discovering a personalized conversation across the different chapters of the game.

Game Design

The gameplay is developed according to the following structure: firstly, the player unlocks a conversation through clickable spheres (StarObjects) following a series of blue watercolor scenes making several choices corresponding to several constellations drawn. Secondly, the NPC acts as a companion and is able to respond to the player depending on the player's choice, using similar mechanics as the player does.



Capture of Mempathy video game. Both the player and the NPC unlock a conversation clicking on the Stars represented as spheres that unlock a conversation in between them.

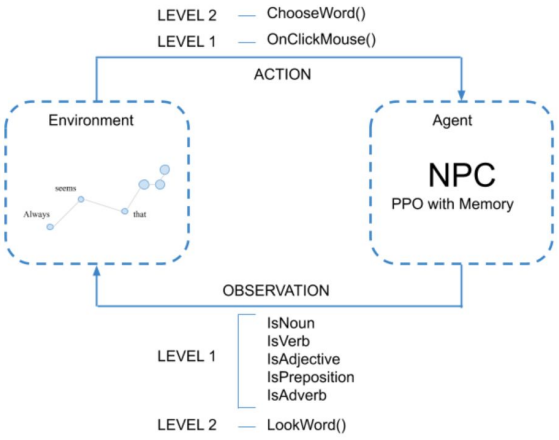
The NPC develops itself under two principles that help the development of the character through the game and its interaction with the player: as the first principle, the one of personhood, defined as the overall impression that the NPC is an independent person, is reflected in the video game by the NPC having its motivations towards the player (offer encouragement, acceptance, and empathy), with the presence of animated eyes inside the game, and using the same gameplay of the player for guiding the conversation. The second principle is bonding: as shared experiences build a deep sense of connection, one of the game's main objectives is to create a bound between the player and the NPC. One of the key challenges here is to overcome some of the factors that could entail a lower bounding, such as superficial and incoherent response or repetitive dialogue. Therefore, the right choice of the

machine learning techniques in this area has been key, as reinforcement learning techniques are oriented towards a specific goal that serves as a motivation for the NPC from the game design perspective.

Reinforcement Learning Environment

Mempathy is a discrete partially observable environment: at each episode, the agent clicks on a series of game objects called StarObjects. Each StarObject has a property attached to the game object corresponding with the word's grammatical structure. Each grammatical structure is connected to a database that contains a list of words.The episode terminates when the agent has clicked in all the stars.

Mempathy Reinforcement Learning environment is constructed at two levels in both the action and observation space: on the observation space, the agent observes the word's grammatical structure (isAdjective) and looks for the word inside the category that has been chosen (happy). This produces an observation vector of dimension 1x6. From the action space, the agent clicks on the object and selects a word based on a probability distribution from the previous one.



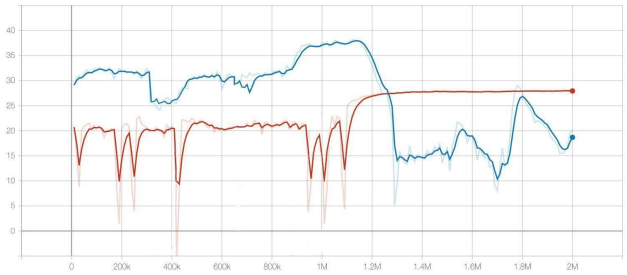
In the Reinforcement Learning Environment, at each time step t, the agent observes the word's grammatical structure (eg: a noun, adjective, verb, adverb) and looks to the word that corresponds to the given structure. It chooses a word from the database and takes the action of clicking in the star.

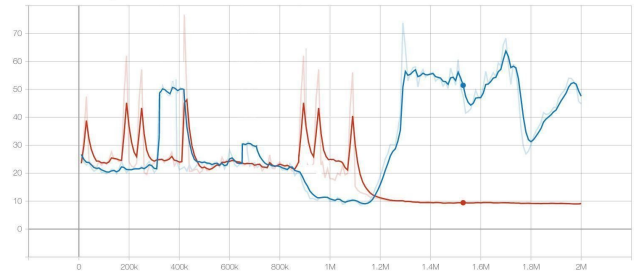
Reinforcement Learning and Imitation Learning

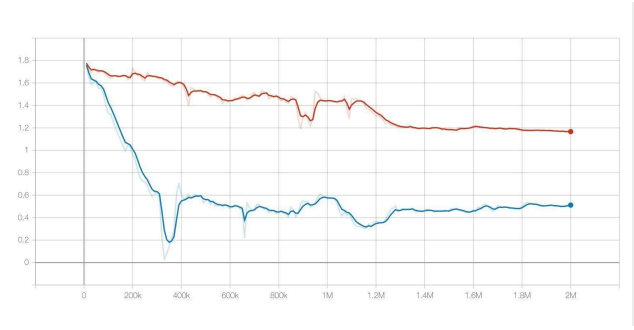
From a general overview, Reinforcement Learning is a method based on learning towards a goal and Imitation learning is a method based on learning from expert demonstration. Both are methods for sequential tasks, where the agent comes up with a policy in order to achieve the optimal performance. The difference, however, is that in Imitation learning, the agent first observes the actions of an expert during the training phase. The agent uses this training set to learn a policy that tries to mimic the actions demonstrated by the expert, in order to achieve the best performance. In Reinforcement Learning there is no such expert and the agent has a reward function, and it explores the action space for coming up by itself (using trial and error) with an optimal policy.

Results and training

As shown in the figures, the Imitation Learning experiment shows more stable results in a 2 million steps training as it shows that rewards go up, while episode length goes down. Entropy measurement also shows less randomness in the information being processed.







PPO (blue) and GAIL (red) Comparison of cumulative reward, episode length, and policy loss

Conclusions and future developments

With respect to the game design perspective, the principles of personhood, bonding, and value can offer key hints for designing NPCs under the goal of creating companionship.

Regarding environment design, language can offer an opportunity to use Reinforcement Learning techniques that align with the Reinforcement Learning challenges such as large space complexity and sequence dependence problem. With respect to reward design inside Reinforcement Learning, the future stands to design directly towards a type of emotion-driven by motivation coming from the NPC player.

Imitation learning shows faster activatable results and desirable and controlled behavior during play. If we want a fully controlled experience, Imitation Learning can significantly speed up video game's construction. Besides, using Imitation Learning, the video game industry could introduce in future developments Human in the Loop techniques or players as teachers, designing more personalized experiences for games.

References

Hiwiller, Z; Sail, F. 2018 *Group Report: Designing Feelings of Companionship with Non-Player Characters* The Thirteenth Annual Game Design Think Tank Project Horseshoe

Juliani, A.; Berges, V.; Teng, E.; Cohen, A.; Harper, J.; Elion, C.; Goy, C.; Gao, Y.; Henry, H.; Mattar, M.; Lange, D. 2020 *Unity: A General Platform for Intelligent Agents.*

Parreño, G. 2020 Mempathy video game playable demo. https://soygema.itch.io/mempathy

Shao, K.; Tang, Z.; Zhu, Y.; Li, N.; Zhao, D.; 2019. *A survey of Deep Reinforcement Learning in Video Games*. IEEE

Sutton, R. S; Barto, B. G. 2014. *Reinforcement Learning: An Introduction.*